

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/57982> holds various files of this Leiden University dissertation.

Author: Radosavljevik, D.

Title: Applying data mining in telecommunications

Issue Date: 2017-12-11

Bibliography

- 3gpp (1999), 'RNSAP Cell Load Information Procedure and Message Contents'. Retrieved from http://www.3gpp.org/ftp/tsg_ran/wg3_iu/TSGR3.07/Docs/Pdfs/R3-99c57.PDF.
- Ahn, J. H., Han, S. P. and Lee, Y. S. (2006), 'Customer Churn Analysis: Churn Determinants and Mediation Effects of Partial Defection in the Korean Mobile Telecommunications Service Industry', *Telecommunications Policy* **30**(10), 552–568.
- Alberts, L. J. S. M. (2006), Churn Prediction in the Mobile Telecommunications Industry, Master's thesis, Department of General Sciences, Maastricht University.
- Anil Kumar, D. and Ravi, V. (2008), 'Predicting Credit Card Customer Churn in Banks Using Data Mining', *International Journal of Data Analysis Techniques and Strategies* **1**(1), 4–28.
- Archaux, C., Martin, A. and Khenchaf, A. (2004), 'An SVM Based Churn Detector in Prepaid Mobile Telephony', *International Conference on Information and Communication Technologies (ICTTA)* pp. 19–23.
- Au, W., Chan, K. and Yao, X. (2003), 'A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction', *IEEE Transactions on Evolutionary Computation* **7**(6), 532–545.
- Backstrom, L., Huttenlocher, D., Kleinberg, J. and Lan, X. (2006), 'Group Formation in Large Social Networks: Membership, Growth and Evolution', *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* pp. 44–54.
- Ballings, M. and Van den Poel, D. (2012), 'The Relevant Length of Customer Event History for Churn Prediction: How long is long enough?', *Expert Systems with Applications* **39**(18), 13517–13522.
- Bermolen, P. and Rossi, D. (2009), 'Support Vector Regression for Link Load Prediction', *Computer Networks* **53**(2), 191–201.

- Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K. and Wiswedel, B. (2007), KNIME: The Konstanz Information Miner, in 'Studies in Classification, Data Analysis and Knowledge Organization (GfKL 2007)', Springer.
- Birke, D. and Swann, G. P. (2006), 'Network Effects and The Choice of Mobile Phone Operator', *Journal of Evolutionary Economics* **16**(1–2), 65–84.
- Borgatti, S. (1994), 'A Quorum of Graph Theoretic Concepts', *Connections* **17**(1), 47–59.
- Buckinx, W. and Van den Poel, D. (2005), 'Customer Base Analysis: Partial Defection of Behaviourally Loyal Clients in a Non-Contractual FMCG retail Setting', *European Journal of Operational Research* **164**(1), 252–268.
- Carbonneau, R., Laframboise, K. and Vahidov, R. (2008), 'Application of Machine Learning Techniques for Supply Chain Demand Forecasting', *European Journal of Operational Research* **184**(3), 1140–1154.
- Caruana, R. (1997), 'Multitask Learning', *Machine Learning* **28**(2), 41–75.
- Christiansen, T. and Torkington, N. (2003), *Perl Cookbook*, 2 edn, O'Reilly, Sebastopol.
- Dasgupta, K., Singh, R., Viswanathan, B., Chakraborty, D., Mukherjea, S. and Nana-vati, A. A. (2008), 'Social Ties and Their Relevance to Churn in Mobile Telecom Networks', *Proceedings of the 11th International Conference on Extending Database Technology* pp. 668–677.
- Datta, P., Masand, B., Mani, D. and Li, B. (2000), 'Automated Cellular Modeling and Prediction on a Large Scale', *Artificial Intelligence Review* **14**, 485–502.
- Dougherty, J., Kohavi, R. and Sahami, M. (1995), Supervised and Unsupervised Discretization of Continuous Features, in A. Prieditis and S. Russell, eds, 'Proceedings of the Twelfth International Conference on Machine Learning', Morgan Kaufmann, San Francisco, CA, pp. 194–202.
- Ericsson (2013), 'Ericsson Performance Management'. Retrieved from <http://www.ericsson.com/ourportfolio/products/performance-management>.
- Feenberg, D. R., Gentry, W. M., Gilroy, D. and Rosen, H. S. (1989), 'Testing the Rationality of State Revenue Forecasts', *Review of Economics and Statistics* **71**, 300–308.
- Feinberg, E. and Genethliou, D. (2010), Load Forecasting, in J. Chow, F. Wu and J. Momoh, eds, 'Applied Mathematics for Restructured Electric Power Systems', pp. 269–285.

- Ferreira, J., Vellasco, M., Pacheco, M. and Barbosa, C. (2004), 'Data Mining Techniques on the Evaluation of Wireless Churn', *ESANN 2004 Proceedings – European Symposium on Artificial Neural Networks* pp. 483–488.
- Fullerton, T. M. (1989), 'A Composite Approach to Forecasting State Government Revenues: Case Study of the Idaho Sales Tax', *International Journal of Forecasting* 5(3), 373–380.
- Geijer Lundin, E., Gunnarsson, F. and Gustafsson, F. (2003), Uplink Load Estimation in WCDMA, in 'IEEE Wireless Communications and Networking Conference', Vol. 3, pp. 1669–1674.
- Grant, G. G. (2003), *ERP & Data Warehousing in Organizations: Issues and Challenges*, IGI Global.
- Gyan, R., Hui, Z., Zhi-Li, Z. and Jean, B. (2012), 'Are Call Detail Records Biased for Sampling Human Mobility?', *ACM SIGMOBILE Mobile Computing and Communications Review* 16(3), 33–44.
- Hadden, J., Tiwari, A., Roy, R. and Ruta, D. (2006), 'Churn Prediction Using Complaints Data', *International Journal of Intelligent Technology* 13, 158–163.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. H. (2009), 'The WEKA Data Mining Software: An Update', *SIGKDD Explorations* 11(1).
- Hamilton, J. D. (1994), *Time Series Analysis*, Vol. 2, Princeton university press Princeton.
- Harell, F. E. J. (2001), *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*, New York: Springer.
- Heikkilä, J. (2002), 'From Supply to Demand Chain Management: Efficiency and Customer Satisfaction', *Journal of Operations Management* 20(6), 747–767.
- Huawei (2013), 'iManager M2000'. Retrieved from <http://www.huawei.com/en/products/oss/mbb-om-product/imanager-m2000/>.
- Hung, S., Yen, D. C. and Wang, H. (2006), 'Applying Data Mining to Telecom Churn Management', *Expert Systems with Applications* 31(3), 515–524.
- Hwang, H., Jung, T. and Suh, E. (2004), 'An LTV Model and Customer Segmentation Based on Customer Value: A Case Study on the Wireless Telecommunication Industry', *Expert Systems with Applications* 26, 181–188.
- Hyndman, R. J. and Khandakar, Y. (2008), 'Automatic Time Series Forecasting: The Forecast Package for R', *Journal of Statistical Software* 26(3), 1–22.
URL: <http://www.jstatsoft.org/article/view/v027i03>

- IBM (2017), '10 Key Marketing Trends for 2017'. Retrieved from <https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=WRL12345USEN>.
- Idris, A., Khan, A. and Lee, Y. S. (2013), 'Intelligent Churn Prediction in Telecom: Employing mRMR Feature Selection and RotBoost Based Ensemble Classification', *Applied Intelligence* **39**(3), 659–672.
- Kass, G. V. (1980), 'An Exploratory Technique for Investigating Large Quantities of Categorical Data', *Applied Statistics* pp. 119–127.
- Kawale, J., Pal, A. and Srivastava, J. (2009), 'Churn Prediction in MMORPGs: A Social Influence Based Approach', *Proceedings of the 2009 International Conference on Computational Science and Engineering* **4**, 423–428.
- Kendall, M. (1938), 'A New Measure of Rank Correlation', *Biometrika* **30**(1-2), 81–89.
- Kim, H. and Yoon, C. (2004), 'Determinants of Subscriber Churn and Customer Loyalty in the Korean Mobile Telephony Market', *Telecommunications Policy* **28**(9–10), 751–765.
- Koch, R. (2011), *The 80/20 Principle: The Secret to Achieving More with Less*, Crown Business.
- Kohavi, R. and John, G. (1997), 'Wrappers for Feature Subset Selection', *Artificial Intelligence* **97**, 273–324.
- Kraljevic, G. and Gotovac, S. (2010), 'Modeling Data Mining Applications for Prediction of Prepaid Churn in Telecommunication Services', *Automatika* **51**(3), 375–383.
- Kusuma, P. D. (2013), Extending Traditional Telecom Churn Prediction Using Social Network Data, Master's thesis, Leiden University. Retrieved from <http://www.liacs.nl/assets/2013-03PalupiKusuma.pdf>.
- Kusuma, P. D., Radosavljevik, D., Takes, F. W. and van der Putten, P. (2013), Combining Customer Attribute and Social Network Mining for Prepaid Mobile Churn Prediction, in 'Proc. the 23rd Annual Belgian Dutch Conference on Machine Learning (BENELEARN)', pp. 50–58.
- Lemmens, A. and Croux, C. (2006), 'Bagging and Boosting Classification Trees to Predict Churn', *Journal of Marketing Research* **43**(2), 276–286.
- Li, J. and O'Donoghue, C. (2013), 'A Survey of Dynamic Microsimulation Models: Uses, Model Structure and Methodology', *International Journal of Microsimulation* **6**(2), 3–55.
- Lima, E., Mues, C. and Baesens, B. (2009), 'Domain Knowledge Integration in Data Mining Using Decision Tables: Case Studies in Churn Prediction', *Journal of the Operational Research Society* **60**, 1096–1106.

- Mäder, A. and Staehle, D. (2004), Analytic Modeling of the WCDMA Downlink Capacity in Multi-Service Environments, in '16th ITC Specialist Seminar', pp. 229–238.
- Malhotra, A. and Malhotra, C. K. (2013), 'Exploring Switching Behavior of US Mobile Service Customers', *Journal of Services Marketing* 27(1), 13–24.
- Marr, B. (2015), 'Big Data: 20 Mind-Boggling Facts Everyone Must Read', Forbes. Retrieved from <https://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read>.
- Meyer, C. and Schwager, A. (2007), 'Understanding Customer Experience', *Harvard Business Review* 85(2), 116–126.
- Microsoft Corporation (2010), 'Microsoft Excel'. Retrieved from <http://office.microsoft.com/en-us/excel/>.
- Miller Jr, R. G. (2011), *Survival Analysis*, Vol. 66, John Wiley & Sons.
- Min, D. and Wan, L. (2009), 'Switching Factors of Mobile Customers in Korea', *Journal of Service Science* 1(1), 105–120.
- Motahari, S., Mengshoel, O. J., Reuther, P., Appala, S., Zoia, L. and Shah, J. (2012), 'The Impact of Social Affinity on Phone Calling Patterns: Categorizing Social Ties from Call Data Records', In *Proceedings of the 6th SNA KDD Workshop* pp. 9–17.
- Mozer, M., Wolniewicz, R., Johnson, E. and Kaushansky, H. (1999), 'Churn Reduction in the Wireless Industry', *Proceedings of the Neural Information Systems Conference*.
- Muckenheim, J. and Bernhard, U. (2001), A Framework for Load Control in 3rd Generation CDMA Networks, in 'In Proc of the IEEE Global Telecommunications Conference', Vol. 6, pp. 3738–3742.
- MyCom (2013), 'NIMS-PrOptima Service & Network Performance Management'. Retrieved from <http://www.mycom-int.com/products/nims-proptima-service-and-network-performance-solution/>.
- Nanavati, A. A., Gurumurthy, S., Das, G., Chakraborty, D., Dasgupta, K., Mukherjea, S. and Joshi, A. (2006), On the Structural Properties of Massive Telecom Call Graphs: Findings and Implications, in 'Proceedings of the 15th ACM International Conference on Information and Knowledge Management', ACM, pp. 435–444.
- Natalizio, E., Marano, S. and Molinaro, A. (2005), Packet Scheduling Algorithms for Providing QoS on UMTS Downlink Shared Channels, in 'IEEE VTC', Vol. 4, pp. 2597–2601.
- Neo4j (2012), 'Neo4j: Community Edition (Version 1.8.M05) [Software]'. Retrieved from <http://Neo4j.org/>.

- Neslin, S., Gupta, S., Kamakura, W., Lu, J. and Mason, C. (2006), 'Detection Defection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models', *Journal of Marketing Research* 43(2), 204–211.
- Ngonmang, B., Viennet, E. and Tchunte, M. (2012), Churn Prediction in a Real Online Social Network Using Local Community Analysis, in 'Advances in Social Networks Analysis and Mining (ASONAM)', IEEE, pp. 282–288.
- Nokia Siemens Networks (2008), 'Rel. RU10- System Library, v.1: RNC Counters - RNW Part'. Proprietary and Confidential.
- Oentaryo, R. J., Lim, E. P., Lo, D., Zhu, F. D. and Prasetyo, P. (2012), 'Collective Churn Prediction in Social Networks', In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* pp. 210–214.
- Oracle (2011), 'Oracle Database Documentation Library'. Retrieved from <http://www.oracle.com/pls/db102/homepage>.
- Pareto, V. (1964), *Cours d'Économie Politique*, Vol. 1, Librairie Droz.
- Pegasystems (2008), 'Predictive Analytics Director (Version CDM 6.3) [Software]'. Retrieved from <http://www.pega.com/products/decision-management>.
- Perner, P. (2002), *Data Mining on Multimedia Data, LNCS, vol. 2558*, Springer Verlag, Berlin.
- Pine, B. J. I. and Gilmore, J. H. (1999), *The Experience Economy*, Harvard Business School Press, Boston, MA.
- Polepally, A. and Mohan, S. (2012), Behavior Analysis of Telecom Data Using Social Networks Analysis, in 'Behavior Computing', Springer London, pp. 291–303.
- Prasad, U. D. and Madhavi, S. (2012), 'Prediction of Churn Behavior of Bank Customers Using Data Mining Tools', *Business Intelligence Journal* 5(1), 96–101.
- QlikTech International AB (2014), 'QlikView Personal Edition (Version 11.2)'. Retrieved from <http://us-d.demo.qlik.com/download>.
- R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
URL: <http://www.R-project.org>
- Radosavljevik, D. and van der Putten, P. (2013), Preventing Churn in Telecommunications: The Forgotten Network, in 'International Symposium on Intelligent Data Analysis', Springer, pp. 357–368.

- Radosavljevik, D. and van der Putten, P. (2014), Large Scale Predictive Modeling for Micro-Simulation of 3G Air Interface Load, *in* 'Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining', ACM, pp. 1620–1629.
- Radosavljevik, D. and van der Putten, P. (2017), Service Revenue Forecasting in Telecommunications: A Data Science Approach, *in* W. Duivesteyn, M. Pechenizkiy, G. Fletcher, V. Menkovski, E. Postma, J. Vanschoren and P. van der Putten, eds, 'Benelearn 2017: Proceedings of the Twenty-Sixth Benelux Conference on Machine Learning', pp. 187–189.
- Radosavljevik, D., van der Putten, P. and Kyllesbech Larsen, K. (2010a), 'The Impact of Experimental Setup in Prepaid Churn Prediction for Mobile Telecommunications: What to Predict, for Whom and Does the Customer Experience Matter?', *Transactions on Machine Learning and Data Mining* 3(2), 80–99.
- Radosavljevik, D., van der Putten, P. and Kyllesbech Larsen, K. (2010b), The Impact of Experimental Setup on Prepaid Churn Modeling: Data, Population and Outcome Definition, *in* I. Bichindaritz, P. Perner and G. Russ, eds, 'Advances in Data Mining, Workshop Proceedings', IBAI Publishing, Leipzig, pp. 14–27.
- Radosavljevik, D., van der Putten, P. and Kyllesbech Larsen, K. (2012), Mass Scale Modeling and Simulation of the Air-Interface Load in 3G Radio Access Networks, *in* 'Advances in Intelligent Data Analysis XI', Springer Berlin Heidelberg, pp. 301–312.
- Richter, Y., Yom-Tov, E. and Slonim, N. (2010), 'Predicting Customer Churn in Mobile Networks Through Analysis of Social Groups', *Proceedings of the SIAM International Conference on Data Mining* pp. 732–741.
- Saravanan, M. and Raajaa, G. V. (2012), A Graph-Based Churn Prediction Model for Mobile Telecom Networks, *in* 'Advanced Data Mining and Applications', Springer Berlin Heidelberg, pp. 367–382.
- Schmitt, B. H. (2003), *Customer Experience Management: A Revolutionary Approach to Connecting with Your Customers*, John Wiley and Sons, Hoboken, NJ.
- Seo, D., Ranganathan, C. and Babad, Y. (2008), 'Two-level Model of Customer Retention in the US Mobile Telecommunications Service Market', *Telecommunications Policy* 32(3), 182–196.
- Seshadri, M., Machiraju, S., Sridharan, A., Bolot, J., Faloutsos, C. and Leskovec, J. (2008), 'Mobile Call Graphs: Beyond Power-Law and Lognormal Distributions', *In Proceedings of the 14th ACM SIGKDD* pp. 596–604.
- Smith, S. and Wheeler, J. (2002), *Managing the Customer Experience: Turning Customers into Advocates*, FT Prentice Hall, Harlow, USA.

- Strawberry Perl (2011), 'Strawberry Perl'. Retrieved from <http://strawberryperl.com/>.
- Svoboda, P., Buerger, M. and Rupp, M. (2008), Forecasting of Traffic Load in a Live 3G Packet Switched Core Network, in J. Chow, F. Wu and J. Momoh, eds, 'Proc. of 6th International Symposium on CNSDSP', pp. 433–437.
- Teradata Corporation (2017), 'Teradata Online Library'. Retrieved from http://info.teradata.com/HTMLPubs/DB_TTU_15_00/index.html.
- Trapletti, A. and Hornik, K. (2017), *Tseries: Time Series Analysis and Computational Finance*. R package version 0.10-40.
URL: <https://CRAN.R-project.org/package=tseries>
- Trueman, B., Wong, M. F. and Zhang, X.-J. (2001), 'Back to Basics: Forecasting the Revenues of Internet Firms', *Review of Accounting Studies* **6(2–3)**, 305–329.
- Tsamardinos, I. and Aliferis, C. (2003), 'Towards Principled Feature Selection: Relevancy, Filters and Wrappers', *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics* .
- Urbanek, S. (2014), *RJDBC: Provides Access to Databases Through the JDBC Interface*. R package version 0.2-5.
URL: <http://CRAN.R-project.org/package=RJDBC>
- van der Putten, P. and van Someren, M. (2004), 'A Bias-Variance Analysis of a Real World Learning Problem: The CoIL Challenge 2000', *Machine Learning* **57(1-2)**, 177–195.
- Veeramachaneni, K. (2016), 'Why You're not Getting Value from Your Data Science', Harvard Business Review. Retrieved from <https://hbr.org/2016/12/why-youre-not-getting-value-from-your-data-science>.
- Verbeke, W., Martens, D., Mues, C. and Baesens, B. (2011), 'Building Comprehensible Customer Churn Prediction Models with Advanced Rule Induction Techniques', *Expert Systems with Applications* **38(3)**, 2354–2364.
- Wang, Y., Cong, G., Song, G. and Xie, K. (2010), 'Community-based Greedy Algorithm for Mining Top-K Influential Nodes in Mobile Social Networks', *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* pp. 1039–1048.
- Weatherford, L. R. and Belobaba, P. P. (2002), 'Revenue Impacts of Fare Input and Demand Forecast Accuracy in Airline Yield Management', *Journal of the Operational Research Society* pp. 811–821.

- Weatherford, L. R. and Kimes, S. E. (2003), 'A Comparison of Forecasting Methods for Hotel Revenue Management', *International Journal of Forecasting* **19**(3), 401–415.
- Wei, C. and Chiu, I. (2002), 'Turning Telecommunications Call Details to Churn Prediction: A Data Mining Approach', *Expert Systems with Applications* **23**, 103–112.
- Wickham, H. and Francois, R. (2016), *Dplyr: A Grammar of Data Manipulation*. R package version 0.5.0.
URL: <http://CRAN.R-project.org/package=dplyr>
- Wirth, R. and Hipp, J. (2000), CRISP-DM: Towards a Standard Process Model for Data Mining, in 'Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining', pp. 29–39.
- Witten, I. H. and Frank, E. (2005), *Data Mining: Practical Machine Learning Tools and Techniques*, 2 edn, Morgan Kaufmann, San Francisco.
- Wolpert, D. H. and Macready, W. G. (1997), 'No Free Lunch Theorems for Optimization', *IEEE Transactions on Evolutionary Computation* **1**(1), 67–82.
- Yates, R. (1995), 'A Framework for Uplink Power Control in Cellular Radio Systems', *IEEE JSAC* **13**(7), 3141–3147.
- Ziegler, C. N. and Lausen, G. (2004), 'Spreading Activation Models for Trust Propagation', In *Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service* pp. 83–97.

